R E P O R T R E S U M E S

ED 003 210 56

USE OF EXPLORATORY RESEARCH AND INDIVIDUAL TUTORING TECHNIQUES FOR THE DEVELOPMENT OF PROGRAMING METHODS AND THEORY, FINAL REPORT.

BY- SILBERMAN, H. AND OTHERS
SYSTEM DEVELOPMENT CORP., SANTA MONICA, CALIF.

REPORT NUMBER NDEA-VIIA-968-3 PUB DATE 29 JUN 64
REPORT NUMBER TM-895/200/00
GRANT OEG-7-14-0000-181
EDRS PRICE MF-$0.18 HC-$2.48 62P.

DESCRIPTORS- *PROGRAMED INSTRUCTION, *EDUCATIONAL RESEARCH, LEARNING DIFFICULTIES, INDIVIDUAL STUDY, *INDIVIDUAL DIFFERENCES, TUTORING, *PROGRAM DESIGN, PROGRAM DEVELOPMENT, *LEARNING EXPERIENCES, READING, ARITHMETIC, GRADE 1, SPANISH, GEOMETRY, HIGH SCHOOL STUDENTS, SANTA MONICA, CALIFORNIA

FOUR SELF-INSTRUCTIONAL PROGRAMS WERE STUDIED AND REVISED BY USE OF TUTORING TECHNIQUES WITH INDIVIDUAL STUDENTS TO CREATE MODIFIED PROGRAMS SUPERIOR TO THE ORIGINAL ONES. THE PROGRAMS USED IN THE STUDY WERE FIRST-GRADE READING, FIRST-GRADE ARITHMETIC, JUNIOR HIGH SCHOOL SPANISH, AND SENIOR HIGH SCHOOL GEOMETRY. PROGRAMS THAT LED TO IMPROVED STUDENT PERFORMANCE WERE THEN COMPARED TO ISOLATE THOSE OPERATIONS COMMON TO ALL FOUR PROGRAMS. SYSTEMATIC METHODOLOGICAL DIFFERENCES FROM ONE CONTENT AREA TO ANOTHER AND FROM ONE AGE LEVEL TO ANOTHER WERE ALSO EXAMINED. THREE GENERAL PRINCIPLES OF PROGRAM IMPROVEMENT WERE ISOLATED--(1) EVERY SKILL SHOULD BE INCLUDED IN THE PROGRAM OBJECTIVES, AND EVERY SUBSKILL THAT LOGICALLY CONTRIBUTES TO THOSE OBJECTIVES SHOULD BE EXPLICITLY COVERED BY THE PROGRAM UNLESS IT EXISTS IN THE STUDENT'S ENTRY REPERTOIRE, (2) ANY MATERIALS THAT DO NOT CONTRIBUTE TO THE PROGRAM OBJECTIVES SHOULD BE ELIMINATED, AND (3) THE STUDENT SHOULD BE REQUIRED TO DEMONSTRATE MASTERY OF EACH COMPONENT SUBSKILL BEFORE HE IS ALLOWED TO ADVANCE TO NEW TOPICS THAT ARE BASED ON THE EARLIER MATERIALS. (WB)

AD603210

963-3

94

TM-895/200/00

Final Report

Use of Exploratory Research and Individual Tutoring Techniques for the Development cf Programming Methods and Theory

29 June 1964

TM-895/200/00

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

ED 030210

TECHNICAL MEMORANDUM

(TM Series)

The research reported herein was conducted under SDC's independent research program and a grant from the Office of Education, U. S. Department of Health, Education, and Welfare.

Final Report

Use of Exploratory Research and Individual Tutoring Techniques for the Development of Programming Methods and Theory

by

H. Silberman & J. Coulson (Principal Investigators)
and R. Melaragno and G. Newmark

SYSTEM DEVELOPMENT CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA

Title VII of P.L. 85-864
National Defense Education Act of 1958
Grant 7-14-0000-181
29 June 1964

The views, conclusions, or recommendations expressed in this document do not necessarily reflect the official views or policies of agencies of the United States Government.

Excepting agencies of the United States Government, permission to quote from this document or to reproduce it, wholly or in part, should be obtained in advance from the System Development Corportion.

© Copyright by System Development Corporation, 1964.

(page 2 blank)

ACKNOWLEDGMENTS

Grateful acknowledgment is given to the following individuals, who made valuable contributions in preparing and administering the programmed materials:

James Caballero
Nancy Echols
Ralph Emch
Evelyn Westrem
Judy Field
Ernest Garcia
Patrick Herrera
Joanne Robinson
Pearl Roosinck
Richard Schutz
Carleton Shay

The authors also wish to thank the principals and teachers in the Santa Monica (California) School District and the Culver City (California) Unified School District for their help in making children available for this research.

PRECEDING PAGE BLANK-NOT FILMED

TABLE OF CONTENTS

Page

INTRODUCTION. 5

EMPIRICAL DEVELOPMENT OF PROGRAMS IN READING, ARITHMETIC, SPANISH, AND GEOMETRY 7

Reading 7
Arithmetic. 11
Spanish 13
Geometry. 18

GENERALIZATIONS FROM EXPLORATORY PHASE. 23

Comparison of Observations. 23
Program Improvement Principles Common to the Four Areas . . . 24

CONCEPTUAL MODELS 27

C. E. Osgood. 28
B. F. Skinner 28
D. P. Ausubel 29
J. S. Bruner. 31
C. L. Hull. 32
W. K. Estes 32

FINAL EXPERIMENT. 34

Problem 34
Method. 34
Results and Discussion. 38

SUMMARY 41

APPENDICES. 43

PRECEDING PAGE BLANK-NOT FILMED

INTRODUCTION

Most of the large number of experimental investigations of programmed instruction have been of the formal, hypothesis-testing variety, based on a priori assumptions about the variables that influence learning in the self-instructional mode. Due to the lack of adequate exploratory research, important program variables have not been identified and significant results have seldom been obtained in the formal experiments. The study reported here is intended to generate useful and testable hypotheses concerning significant variables in programmed instruction, and to provide the beginning of an empirically based conceptual framework that can, in the future, guide a systematic program of research and implementation.

The present study used exploratory research techniques in the experimentation leading up to the conceptualization stage. Four self-instructional programs were intensively studied and revised, using tutoring techniques with individual students to create modified programs significantly superior to the original ones. Operations performed on the different programs that led to improved student performance were then compared to isolate those operations common to all four programs and therefore likely to be common to the design of any instructional material. Systematic methodological differences from one content area to another and from one age level to another were also examined.

To increase the likelihood of the results being generalizable, the four programs were selected to represent both verbal and quantitative skills at both the elementary and secondary-school levels. The four programs used for study were: first-grade reading, first-grade arithmetic, junior high school Spanish, and senior high school geometry. A different experimenter worked with each program, and interaction among the experimenters was kept to a minimum in order to retain as much independence as possible in operations performed on the program.

In the program revisions, the experimenter served as a tutor while presenting the program to one student at a time. He halted the presentation and provided tutorial assistance whenever the student exhibited difficulty. Records were kept of student difficulties and tutorial procedures that were beneficial; tutorial operations that overcame common student problems were added to the program as revisions. When the experimenter felt that the program had been revised sufficiently, a comparison of the original and revised program was made. If the revised version proved to be both statistically and practically superior to the original version the tutorial sessions were ended; if not, the experimenter subjected the program to further study and modification.

After all four programs had been significantly improved, each experimenter prepared detailed statements of the operations he had performed in modifying his program. These operations were then compared to discover both agreements and systematic differences, with the operations applicable to all four programs

forming a set of principles of program improvement. Three general principles of program improvement were isolated, and current theoretical systems of learning were surveyed to determine which, if any, system would have been most likely to predict the empirically derived principles. All theoretical systems examined were compatible with the derived principles, in that the principles could be cast in the language of any of them; it was therefore evident that there would be no gain in adopting any one system and likely to be some loss in specificity of prediction by such an adoption. Consequently, it was decided that the principles of program improvement would generate the most fruitful hypotheses for further experimentation if retained in their present form.

The remainder of this report describes the research with each of the four programs, the general principles isolated in comparing the results of the research, the outcome from the examination of theoretical systems, and an experimental verification of these principles.

EMPIRICAL DEVELOPMENT OF PROGRAMS IN READING, ARITHMETIC, SPANISH AND GEOMETRY

This section describes the empirical development of the four programs. A more detailed discussion of the exploratory research in these areas, including copies of program versions, tests, aids, dialogues, etc., is included in separate SDC Technical Memoranda, i.e.,Reading, TM-895/100/00; Arithmetic, TM-895/101/00; Spanish, TM-895/103/00; Geometry, TM-895/102/00.

READING

The objective of the reading program was to teach the child to sound out and read trigram combinations of the initial consonants f, r, s, m and word endings an, it, at, and in as shown in Table 1.

Four words were reserved as a test of the child's ability to transfer the skill acquired in the program to novel trigram combinations. These trigrams are indicated by x's along the diagonal of the matrix in Table 1.

Table 1

Program Objectives

Initial Consonants	Trigram Endings: an	it	at	in
f	x			
r		x		
s			x	
m				x

Early revisions to the program were primarily concerned with sequencing problems. At first, words were grouped with a common ending, e.g., man, ran. With this procedure, it was found that children were responding only to the variable initial consonant. When later confronted with an item such as rin they would say ran. The solution to this problem appeared to be a matter of holding the initial consonants constant and varying the word endings. The program was accordingly changed so that words were grouped by common initial consonants. This procedure resulted in greatly increased error rates, accompanied by an annoying tendency of the children to avoid looking at the words.

A series of discrimination exercises was then inserted. Trigram pairs were presented which differed, in order of increasing difficulty, in all three letters, in the first letter, in the last two letters, in the last letter, and finally in the middle letter. This procedure somewhat reduced the number of errors in the program, but did not alter criterion performance at all. The next model featured a cumulative sequencing procedure in which each new word was contrasted with all preceding words before the next word was presented.

In addition, an assortment of other techniques was used to improve the program further before undertaking a formal experimental comparison with the initial version. Item formats were simplified; pictures and mnemonics were used to evoke vowel sounds, e.g., Indian for i, apple for a; stories were designed around the trigrams, e.g., "Mr. Ban is a man with a fan"; verbal reminders were issued such as, "remember the letter a goes with the /ae/ sound"; and so on.

The comparison of the revised version and the original version yielded small differences on program words and no difference on transfer words. The assumption that children would quickly induce the phonic generalizations without explicit development of these subskills within the program was abandoned.

The program was redesigned using a synthetic approach that began with a single letter as the unit and proceeded to the whole trigram in a fashion similar to a chaining procedure described by Jim Evans at Teaching Machines Incorporated. Part of the new program taught children to phoneticize trigrams. Another part was designed to teach children to blend or amalgamate the sound elements. The sound elements produced by the child when "sounding out" the trigrams were supposed to provide the occasion for his saying the whole trigram. Instead, the phoneticizing response and the whole trigram response appeared to be separately and independently associated with the grapheme; the phoneticization did not evoke the whole trigram pronunciation. Consequently, children had no success on the transfer words even though they could read the program words.

Various solutions to this problem were tried. Breaking the trigram into three elements rather than two merely intensified the problem. Making the partition between the vowel and final consonant rather than between the initial consonant and the vowel was also ineffective.

Other techniques that were tried included preliminary practice on amalgamating syllables, and a gradual increase in the time interval between audio segments of phoneticized trigram stimuli. Another technique given a short run was the use of 5x8 cards whose ends were folded over the inscribed grapheme. Children were to sound out the elements as they unfolded left and right flaps respectively. The rate of this manipulation was increased until the whole word was exposed at once. After several other methods were tried, a simple combination of an echoic and fading procedure was found most effective in teaching children to amalgamate the sound elements. For example, at first /m/ /an/ /man/ was echoed rhythmically by the child. Later he responded to /m/ /an/ with /m/ /an/ /man/. It was found

particularly important here not to allow the child to continue with the program until this segment had been completely mastered. When children faltered on this segment, they were branched back and given extended practice until their pronunciation was brought under control of their own phoneticization.

At this point, children taking the program were still unable to cope with the transfer words. Special practice in making the transfer to novel trigram combinations within the program had to be given before they were able to decode the novel combinations on the criterion test. This practice was accomplished by omitting the feedback stimulus following four selected words within the program. If the child was unable to sound out and read these words, he was branched to familiar review words that contained the elements of the word giving him trouble. If he still could not read the novel word, he was branched to other familiar words that rhymed with that word. Then he was once again led through the sounding out procedure step by step before he was given the correct pronunciation. This procedure was repeated for each of the four words selected.

Other changes were made to improve the program before making another formal comparison with the original version. These consisted primarily of removing those segments of the program that required irrelevant responses and were more of a distraction than a help to the child. It was interesting to observe how an absolutely essential ingredient in a previous version became irrelevant in the current one.

In the course of developing the final program, it was observed that taped versions did not do as well as the identical versions presented orally by an experimenter. One hypothesis for explaining this phenomenon is that the experimenter modifies the pacing and the prosodic features of his language in response to subtle cues provided by the learner. In other words, the experimenter presenting the program is simultaneously making frequent diagnostic assessments of the student's behavior. For example, if the student is smiling and giving quick loud responses, the pace is increased; if he is frowning and hesitant, even though answering correctly, the pace is reduced and discriminative cues are provided to the learner by the experimenter changing the inflection of his words and giving small unintentional gestural responses. Thus, even if the experimenter uses the same words that have been recorded on tape, a certain responsiveness is lost in the instrumentation of the program with a consequent reduction in performance. An alternative hypothesis attributes this effect primarily to a history of social conditioning that endows a human teacher with certain reinforcing properties not shared by the equipment. Further research is needed to test such hypotheses.

The final version of the program contains eight segments. The first segment requires the learner to make an auditory discrimination between initial consonant sounds. In the second and third segments, common words are presented and children must press a key on which is inscribed the initial consonant or the vowel-consonant ending contained in the word. The fourth and fifth segments contain a series of discrimination items that require students to match a visual element to an auditory sample. The sixth and seventh segments teach pupils to amalgamate the sound

elements and require them to use skills learned in earlier parts of the program. The final segment gives them practice in transferring the reading skill to novel words.

The latest version of the program was adapted for use in CLASS* so that a final comparison with the initial version could be made. A group of 30 pupils from three B-1 classrooms was assigned to two treatment groups. Groups were equated on reading-readiness pretest scores. The initial version of the program was given to one group; the final version to the other. The initial version required eight 40-minute sessions; the final version required an average of eleven 40-minute sessions. The final version used branching and included more items than the initial linear program.

The results of the comparison showed that the final group read significantly more words than the initial group on the criterion test. The greatest difference was found on the transfer words. On the average, the final group read three of the four transfer words while none of the initial group read any transfer words.

Observations that were made during the reading research yielded the following generalizations:

1. Stimulus variation. It is necessary to vary the stimuli to which discriminative responses are to be learned so that the students behavior does not come under the control of inappropriate stimulus dimensions.

2. Gaps. All subskills comprising the educational objective must be explicitly developed within the program.

3. Irrelevancies. Material requiring responses that are irrelevant to the criterion task should be eliminated from the program.

4. Mastery. Children should not be permitted to continue through the program unless they have given the desired behavior at each step in the program.

---

*CLASS is a highly flexible computer-based research laboratory with 20 independent study stations that provide both audio and visual presentation on an individualized basis. The computer automatically directs each student to the correct part of the program, analyzes and evaluates all student responses inserted through pushbuttons, and provides individualized feedback to the student about the accuracy of his response. It also records the program sequence for each student and the complete record of student responses, and can generate up-to-date visual displays of past and present student performance for use by the experimenter.

ARITHMETIC

The arithmetic program was developed from an approach to modern mathematics, at the elementary school level, based on set theory (Patrick Suppes, Sets and Numbers, Book 1A, New York: Blaisdell Publishing Co., 1962). Book 1A is the first in a series of integrated workbooks and teacher's manuals that cover all the arithmetical subject matter typically taught in the elementary grades.

In adapting Book 1A to programmed instruction, its first five topics were selected: (1) how a set is written; (2) equality of sets containing identical elements; (3) equality of sets with Arabic numerals; (4) equality of sets with identical cardinalities; and (5) union of sets.

The Original Program

The self-instructional program was composed of six parts, each of which was covered on one day of training. Two types of items were prepared for each part. "Instructional items," drawn from the teacher's manual, introduced the material to the student; "drill items," from the workbook, provided practice on the skills involved. Instructional items were presented in pictorial form with accompanying verbal (tape-recorded) messages. Feedback to drill items was first provided by way of recorded messages, but was later made a part of the items in visual form. All drill items were in the multiple-choice form; the student selected his answer by placing an X in a box alongside his choice. The original version contained 189 items, 90 instructional and 99 drill.

To measure prerequisite skills a pretest was prepared that covered recognition of letters, recognition of numerals, and ability to count. A posttest was also prepared that contained 18 questions similar to items in the program and 17 questions requiring novel responses representing transfer tasks.

Revision 1

Following tutorial sessions with eight first-grade students, Revision 1 of the program was prepared. Two major changes were made to the program--the addition of construction (manipulation) items and the revamping of the section of the program devoted to review.

The students used in the tutorial sessions frequently had difficulty understanding the concepts treated in the program. To provide a more meaningful and concrete way of covering these concepts, construction items were prepared which enabled the student to work with the concepts using a desk-sized flannel board, cardboard symbols, and plastic animals. Twenty-five construction items were added to the program.

For most students, the most difficult task was the discrimination between sets whose elements were important and sets whose cardinalities were important.

A review section in the original version covered this discrimination, but had to be altered materially in Revision 1 to provide more coverage of the discrimination and less review of the rest of the program.

Revision 1 contained 190 items (one more than the original version) and included 101 instructional items and 89 drill items.

An experimental comparison of the original version and Revision 1 was carried out using 15 students. Revision 1 led to posttest scores which were statistically superior to those produced by the original version, but the difference was not practically significant. In addition, Revision 1 took significantly more time than did the original version.

Revision 2

An analysis of the results of the comparison between the original program and Revision 1 showed that the better students learned satisfactorily from the revised program but the less able students did not. Subsequent tutorial revisions used students who were selected specifically because they were slow learners.

Dictated by two shortcomings in Revision 1--the failure of the slower students to learn from the program and the excessive time taken to complete it--four major changes were made to the program.

First, seven construction items devoted to the discrimination between sets whose elements or cardinalities were important were eliminated, since they did not teach the discrimination.

Second, the program was administered to students individually rather than in a group mode. This was done to allow the better students to progress more rapidly and the slower students to avoid the pressure of keeping up with their more able peers.

Third, special remedial branches were added to the program for students having difficulties with concepts, and forward branches were provided for students exhibiting adequate understanding of concepts. Revision 2 contained 19 remedial items and the capability of skipping past 25 items.

Fourth, a new didactic approach to the discrimination problem was incorporated into the program. This approach provided easily learned responses which could mediate between the stimulus set and the alternative responses. For sets whose elements were important, the student was to name the elements, then find another set with the same elements. For sets whose cardinality was important, the student was to count the elements, then find an answer which was the same as the number of objects counted.

Revision 2, incorporating branching, contained a maximum of 223 items and a minimum of 179 items, of which 129 were instructional and 94 were drill.

An experimental comparison of the original program and Revision 2 was performed using 14 students. The Revision 2 group, on the average, saw nine more items, took six and one-half minutes longer to complete the program, and scored six and one-half points higher on the posttest than did the original version group. The difference between mean posttest scores was both statistically and practically significant, while the difference in mean times was not significant.

Discussion

In revising the arithmetic program to create a new version superior to the original program, five changes were made which contributed to increased learning.

(1) Construction items were added, which made the conceptual material more concrete. The original version assumed that first-grade students could learn from pictorial material and accompanying verbal messages, an assumption that was not justified. The construction items were useful in cases where pictorial and verbal material were not sufficient.

(2) Mediating responses were incorporated into the program to help teach the discrimination between two highly similar stimulus-response units. The original version assumed that students could make the discrimination when trained on each stimulus-response element separately, and it was evident that most students could not. The mediators provided a direct and simple means of making the appropriate responses to the similar stimuli.

(3) Branching, both remedial and forward, was built into the program. Remedial branching was useful in guaranteeing that the student had learned the subject matter, and forward branching was useful in eliminating unnecessary material.

(4) The program was individualized, so that each student could progress at an appropriate rate. Individualization provides for adequate but not excessive time on the subject matter.

(5) Throughout all of the tutorial sessions, the program was continually simplified to keep the student from making undesirable responses. Instructional material was simplified to reduce inattention, to direct the student to critical elements of items, and to make the required response clear and simple to produce. Procedures were also simplified, so that students would not have to use substantive material to learn what it was they were to do.

SPANISH

Description of Initial Spanish Program

A programmed course in beginning Spanish, published by a commercial company, was selected for this study. It is designed for use by students without

assistance of a teacher and attempts to teach simultaneously the four basic language skills; listening comprehension, speaking, reading, and writing. Unit One, the first 316 items of the program, was chosen as the starting point for this study. The unit covers a good sample of the basic Spanish sound system, a limited number of lexical items, and some important aspects of Spanish grammar (e.g., declarative, interrogative, and negative word order in the present tense; position of descriptive adjectives; possessive adjectives; definite and indefinite masculine articles). The program is linear; all students go through all the same items.

The materials consist of a programmed textbook containing instructions to the student, visual stimuli (pictures and writing), space for written responses, and written confirmation of written responses. The student has a mask to cover the correct answers until he has made his response. He then uncovers the mask to check his answer with the correct one. The texts are integrated with tape recordings containing auditory stimuli and auditory confirmation of oral responses.

Criterion Tests

The program content was analyzed to determine the emphasis given to each of the four basic language skills. Based on this analysis, a post-training criterion test was constructed covering each of the four skills. Test items covered all the vocabulary and all the major grammatical points in the program.

The speaking test was divided into four parts (i.e., vocabulary recall, pronunciation, grammatical usage, and fluency) in order to achieve greater objectivity and precision in scoring. Each part of the test contained different items and was administered and scored separately. Instructions for administering and scoring speaking tests were continually revised to facilitate objectivity and consistency among judges who evaluated tape recordings of student performance in the final experimental comparison.

Physical Setting

Except for one of the experimental comparisons, which was conducted in the language laboratory of Culver City High School, the study took place in SDC's CLASS facility.

During the individual learning sessions, the students worked at a desk in a small, enclosed room which resembled a recording booth. A one-way glass enables an observer, outside the room, to look in without being observed by the occupant. An overhead microphone, hanging from the ceiling, was attached to a tape recorder outside the room and recorded all student oral responses. The student had the programmed text and a tape recorder in order to play the auditory portion of the program. All instruction took place at SDC after regular school hours and the students were told not to practice elsewhere.

The experimental comparisons, involving 20 students simultaneously, took place in CLASS's large instructional area, which was divided into two ten-student classrooms.

Use of Tutorial Techniques to Revise Programmed Materials

Twenty-nine seventh-grade students took part in individual tutoring sessions. Notes were kept on remedies tried and the results. When a particular remedy appeared effective, an attempt was made to incorporate it into subsequent sessions with the same student and with later students. Such revisions to the program were made, as needed, on a day-to-day basis by writing changes in the programmed text, by inserting new pages in the text, and by replacing portions of the tape-recorded auditory stimuli. When a number of changes made it desirable, complete sections of the text were revised and reproduced and a new master tape recorded. Revisions were based on (1) observations during the tutorial sessions, (2) analysis of written responses in the programmed text, (3) analysis of diagnostic test results, and (4) analysis of the post-training criterion test results. Changes to the original program resulted in five revised versions.

Changes Prior to Version 5

Except for a limited form of branching introduced in Version 4, revisions prior to Version 5 were made within the context of the linear sequence of items.

When students were presented with items contrasting incorrect English sounds and correct Spanish sounds and asked to imitate only the Spanish, they frequently repeated the English or both the English and the Spanish. Consequently these items were eliminated from the program.

Instead of introducing new Spanish words by dividing them into syllables for initial practice on each syllable, new words were introduced as a unit.

Originally the student would hear a new word once and immediately imitate it, then hear it a second time and repeat it again. This was revised so that a student would listen to new material three times before speaking. Subsequently he would repeat three times after the Spanish model.

Items were added to the program in order to provide more practice on difficult structures. A much slower build-up in task complexity was provided, especially in regard to writing in Spanish, in which student performance was consistently lowest.

Students had great difficulty in following program directions. Steps were taken to reduce the excessive variability from item to item in required response behavior, which was a major source of the difficulty. Other steps taken included simplifying English instructions, underlining key words, replacing printed instructions with symbols, and presenting directions on tape immediately prior to presentation of stimulus material.

The practice of tutoring a student every time he made an error was discontinued. Short diagnostic tests in each of the four basic language skills were administered at 50-item intervals as a guide to necessary tutoring.

Development of the Final Program Version (Version 5)

Although there were indications that initial revisions were resulting in improvement, differences in achievement on criterion tests were not statistically significant.

The final program version emphasized branching. Separate tracks were provided for slow and fast students. These tracks were parallel in content and format but differed in such dimensions as step size, amount and type of practice, cueing, and sequencing of subskills. Track A was the fast track; Track B, the slow. The same end-of-unit diagnostic tests (one for each of the four basic skills) were used for both tracks. Branching took place at the end of each unit, on the basis of the diagnostic test results. Branching was inserted by individual language skill so that a student who had trouble with one skill undertook remedial exercises in that skill only. Two levels of branching were provided. The first level was a fast branch providing some additional practice for students making minor or careless errors. The second level was a slow branch providing extensive remedial practice for students having serious difficulty with particular linguistic problems.

When a student failed to show improvement after undergoing all of the remedial exercises, he was branched to the teacher for tutoring. Some students were branched into a limited program where instruction was restricted to one or two skills at a time. Figure 1 shows the branching structure of the final program version.

Experimental Comparisons of Original and Revised Programs

First Comparison. The first comparison between the original program and the revised program was in the group mode, with pacing under experimenter control and all students progressing through the program simultaneously.

Twenty seventh-grade students from Culver City Junior High School were randomly assigned to each of two groups (i.e., group one--original program; group two--revised program). The two groups did not differ significantly on the total combined score or on any of the individual tests (i.e., listening, reading, writing, and speaking).

Second Comparison. Twenty seventh-grade students from Culver City Junior High School were randomly assigned to two groups. Each student worked independently, at his own rate of speed, and upon completion of the program was administered an individual test in listening comprehension, reading, writing, and speaking.

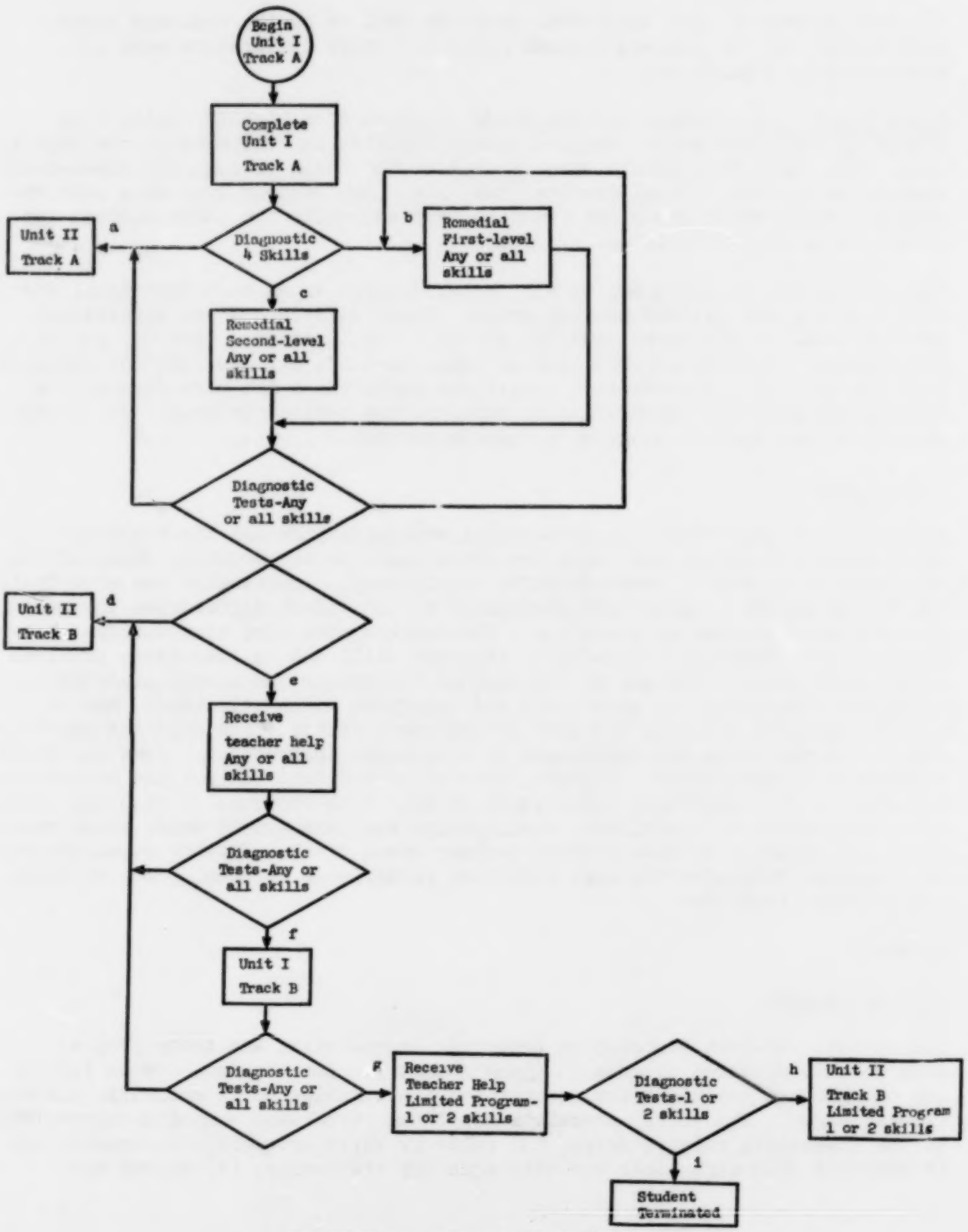

Fig. 1. Branching Structure in Final Program Version

The mean scores on each individual test--as well as on the combined tests--were higher for the revised program group but these differences were not statistically significant.

Final Comparison. Twenty seventh-grade students from Lincoln Junior High School in the Santa Monica Unified School District participated in the experiment. Ten pairs of students were matched on the basis of I.Q. and grade-point average in English, Social Studies, and Math. One student from each pair was randomly assigned to group one (revised program) while the other student was placed in group two (original program).

The mean scores on the total of all tests, as well as on each individual test, were high for the revised program group. These differences are significant for the total of all tests ($t=3.19$, $p<.01$), for listening; ($t=2.07$, $p<.05$), for reading; ($t=2.64$, $p<.01$), for writing; ($t=4.17$, $p<.01$); and for speaking ($t=2.81$, $p<.01$). A one-tailed t-test was used, since the direction of the results was predicted beforehand in favor of the revised program. The groups did not differ significantly on a Spanish pretest.

Conclusions

A significant improvement in achievement was gained through the technique of individual tutoring and iterative revisions. In the process, observations were made which permit some tentative conclusions. Improvement was principally due to making the program more responsive to individual differences in students. This was accomplished by providing a two-track system (for slow and fast learners) and extensive remedial branching by language skill and by linguistic problems within each skill. The use of the teacher to monitor and assist students in following directions, to administer and interpret diagnostic tests, and to provide remedial tutoring was also an important factor. The students generally reacted favorably to the experience of programmed instruction. Some succeeded in working independently. However, most of them seemed to lack the motivation and ability for completely independent study. With students of this age level, some combination of individual, small-group, and large-group study seems necessary, and interaction with a human teacher seems vital. Further experimentation is needed to determine the most effective patterns in the employment of human and material resources.

GEOMETRY

Initial Program

The initial 248-item sequence on geometric inequalities was taken from a commercially produced program on plane geometry. The sequence covers (1) the use of symbolic notations for inequality relationships among geometric elements (lines, angles, etc.), (2) translation between verbal and symbolic expressions of the inequality relationships, (3) rules by which inequality statements can be combined with each other and with equality statements, (4) axioms and

theorems describing inequality relationships in different types of geometric figures, (5) examples of geometric problems that can be solved using the inequality axioms and theorems, and (6) proofs of several inequality theorems.

A fixed-sequence format was used, with items contained in printed booklets. Most items required fill-in responses.

Criterion Test

Program objectives were analyzed and a criterion test was prepared. One set of test items (representing minimal transfer) required answers essentially identical to those elicited in the program; a second set (moderate transfer) required application of theorems and axioms presented in the program to solve geometry problems; and the third set of items (maximal transfer) involved the synthesis of response elements (i.e., subskills) to form proofs of geometric theorems not previously seen by the students.

First Major Program Revision: Elimination of Superfluous Items

Tutorial sessions with the original program suggested that many items were probably superfluous, since they occurred in the sequence after most students had already demonstrated mastery of the topics covered by those items. A shorter form of the program was prepared in which approximately 70 of the original items were eliminated. An experimental comparison between the original and the short program versions showed no difference in criterion performance, so the short form was adopted for subsequent revisions.

Further Revisions Resulting from Tutorial Sessions

In subsequent tutorial sessions with high-school students from second-semester geometry classes, some additional items were eliminated as irrelevant or distracting. In other parts of the program, items were added to teach prerequisite subskills that had not been covered in the original program, and to strengthen the students' responses on other subskills. Major principles of program design and improvement induced from the tutorial sessions are summarized at the conclusion of this discussion of the geometry research.

Experimental Comparison of Original and Revised Programs

After a series of tutorial sessions, a formal experiment was conducted to determine whether the revised program, containing 167 items, was superior to the original 248-item version. Thirty second-semester geometry students received the original program and 30 others from the same classes were given the revised form; all students then received the same criterion test. Performance of the students who received the revised program was significantly higher on the total test, and on the low- and moderate-transfer subtests. No difference was found on the high-transfer portion of the criterion. It was not possible

to obtain a precise measure of study time, but approximate figures indicated that students who received the revised program required less time than those receiving the original form.

Preparation of Proof Booklets

In an attempt to improve the program's ability to teach theorem-proving behavior, the operations involved in theorem proving were further analyzed, and several component subskills were identified. These subskills included: (1) drawing a figure that adequately represents the theorem, (2) identifying the hypothesis ("given") and the conclusion ("prove") of the theorem, (3) identifying the elements (e.g., lines and angles) and the essential contents (e.g., "angle = angle") in any geometric statement, (4) selecting from a catalog of geometric Ideas (axioms, theorems, definitions) the subset of relevant Ideas, (5) applying a geometry Idea to a specific proof by labeling the elements of the Idea to correspond with the figure, and (6) synthesizing all of the previous subskills to produce a logical chain of steps from the "given" to the "prove."

A set of Proof Booklets, designed to supplement the revised program used in the experimental comparison, was written to cover the theorem-proving subskills. In subsequent tutorial sessions it was discovered that the Proof Booklets taught the individual subskills, but could not bring all students to a high level of performance in synthesizing the subskills. At least half the students still could not start with a verbal statement of a novel theorem and generate the complete chain of logical steps connnecting the "given" to the "prove." Further analysis of the students' behavior indicated that the synthesizing skills depended largely on aptitude and prior experience with similar theorems. Since the necessary range of experience could not be provided within the time limits of the study, it was decided to restrict the program objectives and to modify the criterion test accordingly.

Modification of the Criterion Test

New criterion test items were written to test student knowledge on each of the theorem-proving subskills, and to assess student ability to complete a proof by filling in one or two missing steps in a chain of proof statements. Criterion items requiring the formulation of a complete proof to a novel theorem were eliminated.

Revisions to the Proof Booklets

In subsequent tutorial sessions the Proof Booklets were revised to reflect the altered program objectives. The revised materials were found effective in teaching the theorem-proving subskills, bringing students to an 84% performance level on the relevant portions of the modified criterion test. Because the program objectives had been changed, and because the Proof Booklets dealt with subskills not covered in the original program, no formal comparison was

made between the original geometry program and the final revised version. However, the modified criterion test was given to a group of students who had not received any instruction on the inequality theorems; their performance on the theorem-proving subskills was near chance (approximately 19%), indicating that the markedly higher performance of the other students could be attributed to the revised Proof Booklets.

Summary of Programming Principles in Geometry

This portion of the report is a summary of the most important observations made in the geometry study regarding techniques of program design and improvement.

Addition of Preview Materials. The original program contained almost no items telling the student what topic or series of topics he would next be studying. The addition of such preview items contributed significantly to the effectiveness of the revised geometry sequence.

It appeared that the major function of the preview item was to reduce interference among topics, and among steps in a topic, by providing the student with ready-made categories, or labels, for new information. These conceptual labels helped the students to associate each new piece of information with other pieces in the same category, and to make important discriminations between the new piece and other pieces in other categories. A preview covering several topics also helped the student create meta-categories, by telling him what topics (e.g., subskills or specific facts) could be conceptually linked with or subsumed under other topics.

Addition of Review Materials. Review materials were added at several points in the program, particularly at the beginning and end of each lesson. The scope and placement of the reviews was critical, and had to be planned for a specific schedule of program administration. A review sequence that was appropriate for a one-day interval between study sessions was found inadequate for longer intervals.

One function of the review items was to strengthen individual subskills by giving the students extra practice. A second and less-obvious function was to aid the students in organizing a sequence of topics, or a series of steps within a complex topic. In this sense, the review items complemented the preview items in helping students to relate topics to each other and to discriminate among the different topics.

Elimination of Irrelevant and Distracting Materials. Some of the most important improvements in the geometry program resulted from the elimination of items. It was found that some items failed to contribute to the program objectives, and under some conditions actually interfered with learning, either because they were irrelevant or because they provided superfluous over-training on

skills previously acquired by the students. Maximum interference occurred when the irrelevant items required responses that completed with, or were incompatible with, the criterion responses.

When the program was revised, every item was examined to determine whether it contributed toward some specific program objective, and items that did not pass this test were eliminated.

Addition of Materials to Fill Gaps. Probably the most common weakness of the geometry program was the failure to present all of the basic information on a given topic. Gaps resulting from such omissions caused errors on the topics concerned in both the program and the criterion test. Because geometry is hierarchically structured, moreover, the gaps caused errors on subsequent topics that were logically dependent on a complete mastery of the earlier materials.

Some gaps were identified by a logical analysis of the existing items in relation to the program objectives. For each prerequisite subskill identified, the program was examined to see whether there were items that explicitly covered that subskill.

The student's responses in tutorial sessions were the major means of identifying gaps. Errors on a particular topic suggested that some of the subskills relevant to that topic had not been covered.

Addition of Branching Capabilities.

In the geometry program, because of the hierarchical structure of the content area, the student's performance was seriously degraded whenever he advanced from one topic to the next without having mastered the earlier skills. Yet the original program made no provision for ensuring that students of different ability levels received the amount of instruction they required on each topic before moving ahead. In the revised program, branching capabilities were added at several points to adapt the materials to individual needs. One type of branch automatically sent the student back for further review on a topic if his performance on the topic did not meet certain standards. Another type permitted the student to branch back at his own option to any previous topic on which he believed himself to be weak.

GENERALIZATIONS FROM THE EXPLORATORY PHASE

COMPARISON OF OBSERVATIONS

Following the exploratory phase of the research, the four investigators held a series of meetings during which the generalizations and observations that had been made in the four areas were contrasted. The observations were examined to see how they differed (a) for first-grade children as contrasted with secondary school children, (b) for verbal as contrasted with quantitative subject matter, (c) for school children with high and low scholastic aptitude, and (d) for learning objectives requiring three different degrees of transfer.

Age Level

Differences in the age of subjects produced the most striking differences in experimenters' observations. The following points of difference between programming techniques that were effective for older and younger children were observed.

1. Verbal previews and summaries used to establish relationships among various topics were more effective with older children than with the first-graders. Evidently the older child's more extensive response repertoire established a broader base on which to build analogies and other verbal frameworks that integrate the sequence of responses in the program into a coherent unit. Also, older students seemed better able to retain and apply the frameworks over a lengthy sequence of items.

2. What constituted a transfer test question was markedly different for the two age levels. With the younger children, the degree of departure in the criterion test from what was explicitly covered in the program had to be very small. When test items were slightly altered from the conditions in the program, performance declined markedly. With the older students, a much larger discrepancy between program items and test questions was possible before generalization decrement occurred.

3. Older children were better able to evaluate their own responses. Procedures designed to allow students to determine their own sequence did not work as well with younger children as with older children.

4. The first-graders required smaller steps, more frequent overt responses, and more frequent reinforcement.

5. The older students were better able to deal with items that contained a substantial amount of verbal material. The first-graders became inattentive when even a small amount of verbal direction preceded the task. Prompting them during the performance of a task was a better procedure than telling them about the task.

Content Differences

There was little difference between procedures that were effective with the reading program and the first-grade arithmetic program, nor with the Geometry and Spanish programs. Whether the subject matter was mathematics or language was less important than age level in the determination of effective programming procedures. Some differences in the nature of the skills dictated different stimulus or response modes, but these did not affect the techniques of improving the program.

Aptitude Level

Differences in programming techniques mattered less with higher-aptitude students than with lower-aptitude students. The brighter children seemed to compensate for the deficiencies in a program and to learn almost irrespective of programming techniques. Slow children revealed the effects of a deficient program and did not learn unless the program taught them.

Level of Transfer

In all four programs the transfer skills were more difficult to establish than the non-transfer skills. The transfer objectives required the addition of program segments that produced skills mediating the transfer behavior. Branching students back to a program topic until they had learned the appropriate skills facilitated transfer. Students who lacked component skills could not perform the transfer tasks.

PROGRAM IMPROVEMENT PRINCIPLES COMMON TO THE FOUR AREAS

After comparing observations on program improvement procedures, an attempt was made to define an optimal procedure to be used in program revision. It turned out that little agreement could be found in the specific techniques used in the four subject areas. In fact, there was considerable variation from day to day in approaches to the program improvement task within each of the areas. The experimenters could not provide a simple flow chart of their own activities.

Some agreement was found among the four areas by a process of cross-classification. Each of the investigators prepared a set of cards presenting detailed statements of the operations performed in modifying his own program. The cards were exchanged, sorted into categories, reexchanged, reclassified, and then grouped into superordinate categories. By this process, three principles of program improvement common to all four programs were isolated. As short, descriptive labels, these principles were called the "gap" principle, the "irrelevancies" principle, and the "mastery" principle.

Gap Principle

This principle refers to the explicit inclusion in the program of items insuring that all criterion skills are covered. That is, once a set of objectives and a test have been prepared, the program must include items that cover every skill required in the test and every skill prerequisite to the performance of the criterion skills. Most programs fail to cover either some of the criterion skills or some of the intermediate skills which lead up to the criterion skills; the gap principle asserts that a program is improved by the inclusion of items covering the omitted skills. Three general types of items were found which served to fill gaps in a program.

1. Terminal items identical to or similar to questions on the criterion test, i.e., requiring similar responses and/or containing similar item content. For example, the test for the reading program required the student to transfer his learned skills in reading trigrams to novel trigrams not encountered in the program. Practice in making the transfer to novel trigram combinations had to be supplied within the program before students were able to decode the novel combinations on the criterion test.

2. Items logically related to the criterion test, which a content analysis of the test indicates are intermediate steps to the criterion task. For example, the test for the arithmetic program required the student to discriminate between sets whose elements were important and sets whose cardinalities were important. An intermediate step in making the discrimination required the student to recognize the symbology distinguishing a set whose cardinality was important, and items had to be included which provided training on recognizing the distinguishing symbology.

3. Items presented before, during, and following a training session to show the relationships among skills covered in the session and skills covered during earlier sessions. For example, the test for the geometry program required the student to complete the proofs of theorems; therefore, the relationships between steps in theorem proving had to be tied together by explicit coverage in the program.

Irrelevancies Principle

This principle refers to the elimination of unnecessary or distracting items. Most programs contain items and item sequences which require responses unrelated to those contained in the criterion test, and frequently the responses called for in the program are incompatible with responses on the test. Four types of unnecessary items were found in the programs produced in this study.

1. Items designed to produce skills already acquired by the student, as measured by a pretest. For example, a preliminary tutorial session with the geometry program indicated that the skill of measuring line segments had already been learned by the students, and items intended to produce that skill were eliminated from the program.

2. Items not at all similar to the criterion test in the nature of the responses required and/or in the item content (items having no relation to the criterion task). For example, the Spanish program introduced new words by dividing them into syllables, and required responses to the individual syllables. Since the criterion test dealt with whole words, and students trained on syllables frequently continued to respond in terms of syllables on the test, these items were eliminated.

3. Items not contributing to criterion performance, as determined by empirical trials, even though they may have appeared to be similar to criterion tasks or appeared to be necessary steps in reaching criterion skills. For example, the Spanish program contained items contrasting incorrect English sounds with correct Spanish sounds. While it had appeared that training to avoid incorrect English sounds was necessary, empirical trials indicated that the inclusion of this training tended to lead the student to produce the incorrect sounds rather than avoid them, so these items were eliminated.

4. Items directing the student to give procedural responses that serve as a distraction, take unnecessary time, and are incompatible with the task to be learned. For example, the reading program contained copying items in which the student was required to print trigrams. These items proved to be arduous and distracted the child from the reading task so they were eliminated.

Mastery Principle

The mastery principle refers to the provision for individual differences among learners by an increase in program responsiveness. That is, a program can be improved if it is made more flexible, so that individual differences are accounted for and mastery of each skill by every learner is guaranteed. Two methods of modification for mastery were found.

1. Permit individuals to move at their own rate through a fixed sequence of items. For example, both the Spanish and the arithmetic programs were capable of being administered in a group mode, but individualization of instruction avoided many of the difficulties encountered in group instruction.

2. Provide alternative amounts of practice and/or alternative sequences of items, contingent upon student performance. Each program utilized a different method of guaranteeing mastery of skills:

a. The reading program recycled students through previous sequences when program errors exceeded a criterion.

b. The arithmetic program took students to alternative remedial sequences when program errors exceeded a criterion.

c. The Spanish program took students to alternative remedial sequences when errors on diagnostic tests exceeded a criterion.

d. The geometry program allowed the student to review earlier sequences at the student's option, as well as providing programmed review when the student made too many errors.

Ideally, it would be desirable to identify specific structural features of programs that contribute to their effectiveness. To do so would greatly simplify the evaluation and improvement of programs. Principles of program improvement should refer to manipulatable variables that are readily identified upon inspection of the criterion test and the instructional materials. The three principles identified in this study--gaps, irrelevancies, and mastery--refer to the addition, elimination, and repetition of instructional material rather than to intervening variables. They do not, however, depend solely on a rational analysis of the program and the criterion test. To a considerable extent these variables are response defined. A gap or an irrelevancy is identified partially by the extent to which its addition or removal changed criterion performance in an earlier independent situation. Such changes made to a new program in a subsequent experiment are then predicted to have a similar effect on criterion performance. Similarly, with respect to the mastery principle, a diagnostic test for branching a student to remedial material must sample the criterion behavior. If the diagnostic test is judged to be correlated with the criterion in an earlier situation, then satisfactory performance on the criterion test can be predicted with greater confidence. Much further work is still required to develop procedures for defining gaps, irrelevancies, and mastery in terms of stimulus characteristics of the program and criterion test.

CONCEPTUAL MODELS

The three generalizations described above were induced from empirical observations in four independent subject areas. A number of conceptual models were examined in order to find one that corresponded to these three generalizations. The goal of this activity was to identify a theoretical structure that would help to guide subsequent research on programmed materials. It was anticipated that few, if any, of the current theories could be extrapolated to the three principles, or at least that one theory would be far superior to the others in offering specific guidance to the programmer. In fact, however, each of the principles could be readily translated into the descriptive language of any of the models that were examined, but none of the theories furnished specific guidance for program design. As examples, such diverse sources as Osgood, Skinner, Ausubel, Bruner, Hull, and Estes may be considered. Although these systematists have in most cases not extrapolated their models to the applied self-instructional setting, the following interpretations can be inferred from their various writings.

C. E. OSGOOD

Osgood holds that learning involves the modification of the mediation process elicited by a stimulus pattern, or the modification of the instrumental sequence elicited by a mediator, or both. Mediation processes are modified by the regular association of a sign with a stimulus pattern, so that the sign becomes the stimulus for certain of the reactions previously elicited by the original stimulus pattern. Instrumental sequences are modified by the reinforcement of certain reactions to a sign. Thus, through the contiguity of signs and stimulus patterns the signs may acquire stimulus properties, and through reinforcement certain reactions to signs may come to be elicited regularly by the signs.

Gap Principle

The concept of filling a gap could apply to either, or both, of the modifications necessary for learning. Filling a gap may provide the occasion for the association of a sign with a stimulus pattern, leading to mediational modification, or it may provide the occasion for the reinforcement of certain responses to a sign, leading to instrumental modification. A gap in a program could contribute to the failure of either mediational modification or instrumental modification, or both.

Irrelevancies Principle

Eliminating irrelevancies is also applicable to both types of modification. That is, eliminating irrelevancies could mean that incorrect associations of sign stimulus pattern are reduced, since irrelevant stimuli are no longer available. By this interpretation, irrelevant material would disrupt either or both mediational modification or instrumental modification.

Mastery Principle

In the usual experimental situation with which Osgood is concerned, the learner is given sufficient trials to reach a specified criterion. Assuring mastery could be taken to mean the use of sufficient associations between sign and stimulus pattern to bring about mediational modification or the use of sufficient reinforcements of reactions to a sign to bring about instrumental modification (or both).

B. F. SKINNER

Skinner holds that learning is the reassortment of responses in a complex situation. This is taken to mean that a given stimulus acquires control over an operant response (or pattern of operant responses) through the processes of discrimination and shaping. Discrimination takes place when the operant is reinforced in the presence of the given stimulus and not reinforced in the presence of other stimuli. Shaping takes place when operant responses are

selectively reinforced so that behavior successively approximates the desired behavior. In the applied setting, Skinner maintains that occasions must be provided for the reinforcement of particular operants when a given stimulus is present, and that occasions must be provided for the successive approximation of desired complex behaviors so that the approximations can be reinforced.

Gap Principle

Since the acquisition of control over an operant response by a stimulus requires the reinforcement of the operant in the presence of the stimulus, and many such occasions are necessary to establish control, it can be inferred that filling a gap in a program may provide sufficient occasions for the establishment of control. Filling a gap may be alternatively viewed as providing a missing step in shaping a complex behavior.

Irrelevancies Principle

Skinner stresses that for instruction to be effective only certain operants should be reinforced in the presence of a stimulus, and that the reinforcement of other operants makes it difficult (or impossible) for the stimulus to gain control over the desired response. Irrelevant material in a program produces irrelevant responses that can be reinforced and therefore can come under the control of the stimulus. The acquisition of such "superstitious" behavior can interfere with the stimulus control over the desired operant response.

Mastery Principle

Skinner maintains that a stimulus acquires control over an operant only after the operant is reinforced in the presence of the stimulus on a sufficient number of occasions. Therefore, assuring mastery would mean that numerous occasions must be included in a program. Skinner advocates providing a large number of occasions for reinforcement of the operant in the presence of the stimulus to insure stimulus control over the responses of even the slowest learner.

D. P. AUSUBEL

Ausubel* has written a book on meaningful verbal learning that is quite relevant to programmed instruction, being especially directed toward school learning of subject-matter disciplines.

Gap Principle

By Ausubel's own statement, his theory is limited to "reception" learning, which refers to "the situation where the content of the learning task (what is to be learned) is presented to rather than independently discovered by the learner." (p. 1). The inference is clear, therefore, that any learning

*David P. Ausubel, The Psychology of Meaningful Verbal Learning (New York: Grune and Stratton, 1963).

covered by Ausubel's theory must result from specific coverage of the relevant contents in the instructional materials. Ausubel further supports this interpretation by his warning that, "serious breakdowns in learning can often be attributed to inadvertent omission of a logically essential component unit from the total task or to its inadequate integration with other components" (p. 86).

Ausubel does not specify in operational terms what constitutes "omission of a logically essential component unit" in the instruction. He does, however, discuss in general terms certain types of materials that are important for meaningful learning. He states for example, that learning is degraded when the instruction omits explicit discussion of relationships among different ideas presented (p. 80), or when it fails to provide advance "organizers" designed to help the learner subsume new materials under a conceptual framework (p. 81).

Irrelevancies Principle

Ausubel does not give major consideration to the problem of irrelevant and distracting materials in an instructional sequence, or to methods of identifying and eliminating such materials. In another context, he does discuss the need for time efficiency in the educational process (p. 143), so presumably he would advocate the omission of any items that take up time without contributing to the learning objectives.

An important prerequisite in all meaningful learning, according to Ausubel, is an internalized "cognitive structure," or conceptual framework, that has "clarity, stability, and cohesiveness" (p. 76). This framework helps the learner to organize and subsume new materials. Anything that tends to make the framework less stable or more ambiguous must necessarily impair learning. Although Ausubel does not specify the operational factors that degrade cognitive structure, it might be assumed that irrelevant materials and items that call for responses competing with the criterion task responses would have such negative effects.

Mastery Principle

Ausubel states that, "new material in the sequence should never be introduced until all previous steps are thoroughly mastered. Perhaps the chief pedagogic advantage of the teaching machine lies in its ability to control this crucial variable in sequential learning" (p. 86).

Ausubel does not state operational techniques by which the principle of mastery should be implemented in an actual instructional situation. He appears to believe that the mere use of a teaching machine is sufficient, whereas in fact most teaching machine applications simply ensure that the learner is exposed to one step before going to the next, not that he actually masters the earlier step.

J. S. BRUNER

Bruner does not discuss, in any detailed, operational fashion, the way in which human behavior can be effectively controlled. His discussions of the educational process are expressed at the level of general policy and over-all curriculum design. The point that Bruner makes most strongly is that education can be made to have the most lasting impact by giving students an understanding of the fundamental structure of whatever subjects we choose to teach. He does not, however, specify the educational procedures by which the basic principles, once identified, can most effectively be taught.

In his theoretical formulations, Bruner is primarily concerned with the ways in which people categorize the things and events that they encounter. This concern has led him to focus on the experimental situation in which a person is exposed to a series of stimulus patterns (e.g., drawings of people, symbols, or scenery) and attempts to put each of the patterns into one or more categories. Bruner's emphasis is on what the subject himself does with the materials presented. He views the learner not as a passive receiver whose responses simply reflect his physiological limitations and his previous conditioning history, but rather as an active organizer of every event that he encounters. Little or no guidance is given to the instructor who might wish to prepare and sequence educational materials so as to produce systematic changes in a student's behavior.

Gaps and Mastery Principles

Bruner states that "learning often cannot be translated into a generic form until there has been enough mastery of the specifics of the situation to permit the discovery of lower-order regularities which can then be recombined into higher-order, more generic coding systems."* By extrapolation to the programmed learning situation, this statement might be construed to mean that all the preliminary topics and concepts must be specifically covered in the program, as well as the subsequent, logically dependent topics. It would seem to argue, also, that the learner should be required to show mastery of the earlier topics before he is allowed to move ahead in the program.

Irrelevancies Principle

Bruner points out that a given thing may have attributes that do not help define a concept (i.e., attributes that do not help place the thing in any of the prescribed categories). By analogy to the programmed learning situation, irrelevant items may be considered nondefining (and hence useless) items. That is, they do not help the learner define the concepts that he needs to achieve the program objectives, and by inference they should be eliminated.

*J. S. Bruner, "Going Beyond the Information Given," in Contemporary Approaches to Cognition, (Cambridge: Harvard University Press, 1957).

C. L. HULL

Hull's general reinforcement theory of behavior emphasizes that the stimulus is likely in all cases to be an exceedingly complex compound of events. He states that the reaction which will adapt an organism to a given environmental situation depends, as a rule, not upon any single element of that situation but upon a certain combination of elements. Consequently the following interpretations of the three principles can be made with Hull's theory:

Gap Principle

If a stimulus compound as a whole should have an excitatory potential (the tendency to make a given response to a given stimulus) only slightly greater than the response threshold, and if one or more of the elements in the compound be lacking in a given stimulation, the delivery of the incomplete stimulus compound may evoke no reaction whatever.

Irrelevancies Principle

If the afferent impulse received is different from that which was conditioned to the reaction, the excitatory potential will, by a generalization gradient, be less than would be the case if the originally conditioned stimulus had occurred. The alien stimulus in a conditioned stimulus compound reduces the excitatory potential, and as a consequence the amplitude of the resulting reaction will be reduced, its latency will be increased, and the excitatory potential may even fall below the reaction threshold, in which case no reaction at all will occur. Thus irrelevant stimulations will interfere with the capacity of the learner to perform.

Mastery Principle

The excitatory potential depends on habit strength, which is built up through reinforced practice. Insufficient practice will hence result in low habit strength and low excitatory potential. If some component responses are thus deleted from the larger behavior pattern, subsequent responses will suffer since component responses are greatly altered by virtue of their combination with other responses.

W. K. ESTES

Estes treats the totality of stimulus conditions that may be effective during learning as a mathematical set. A subset of these elements is effective on any trial and is assumed to have been drawn randomly from the population of elements. On any given trial those elements that were sampled are connected to a specific response that is reinforced. Although this theory is based on simplified learning situations and cannot be generalized to complex learning, we can find relevant concepts in the theory which might be extrapolated to the three principles as follows:

Gap Principle

In stimulus-sampling models, the set of stimulus elements available for sampling is restricted by the insertion of a gap in the material. Consequently the opportunity for any response becoming associated with those elements is removed.

Irrelevancies Principle

If irrelevant elements are added to the program, they may become conditioned to responses that should have been associated with the relevant elements. The relevant elements that were previously associated with the correct responses may also become reconditioned to responses that are connected to the irrelevant stimulus elements.

Mastery Principle

Each time the student is exposed to a population of stimulus elements, a certain proportion of those elements may become connected to the learned response. If many of the elements are still not conditioned, additional trials with those unconditioned elements increases the probability of their achieving the conditioned state.

While this list is not exhaustive, it does indicate how readily the three principles can be cast into the descriptive language of existing theoretical systems. Initially, the reason for comparing the principles with existing theories was to select a theory to guide subsequent research. However, even though the various theoretical positions are compatible with the derived principles, they offer less specific guidance to the programmer than do the principles themselves. Perhaps by their very nature as abstractions, theories characteristically do not specify the detailed conditions of instruction. The item sequence, the types of information to be included, the techniques of getting out the appropriate behavior, the phrasing of directions, the length of training sessions, and the many other instructional procedures necessary in teaching a given subject matter are not specified by any theory. Perhaps the main function of a theory is to provide a more efficient language for describing effective instructional procedures after they have been discovered empirically.

The principles themselves are at a lower level of abstraction than any theory that might be induced from them. A set of systematically defined principles, being closer to the data, provides more specific help in the design of improved instructional material than any theory. Although it would be possible to construct a theory on the foundation of the generalizations, such a theory would not afford any more understanding of learning phenomena than the existing theories, and in fact would be further from the data. Before a predictive and operationally useful theory can be constructed, more empirical research must

be conducted and more critical variables identified. Such considerations led to a decision in the present study to use the empirically induced principles in their existing fo to generate hypotheses for the final experiment (described in the next section).

FINAL EXPERIMENT

PROBLEM

In the exploratory phase of this project three variables (corresponding to the gap, irrelevancies, and mastery principles) were observed to be important in each of the four areas of study. To test the principles further, another experiment was designed. Whereas the effects of the three principles were originally noted by making improvements to four programs, the intent of the final experiment was to cross-validate those findings by reversing the process. It was hypothesized that if the same improvements were removed from an effective program there should be a corresponding performance decrement, even though the subject area and the sample of students was markedly different from those which yielded the original findings.

The three variables that were tested were the gap variable, the irrelevancies variable, and the mastery variable. The gap variable refers to the presence or absence of items that cover a skill that is logically related to the criterion task. The irrelevancy variable refers to the presence or absence of items that cover a skill that is unrelated to the criterion task. The mastery variable refers to the provision of alternative amounts of practice on items contingent upon student performance.

It was hypothesized that the effects of the three variables would be cumulative. Thus, it was predicted that learning with a program in which mastery was required and which contained neither irrelevancies nor gaps (good version), would be superior to learning with a version that did not require mastery (linear version); that student learning with the linear version would exceed the learning with a linear version to which irrelevant items were added (irrelevant version); and that learning with the irrelevant version would exceed the learning with an irrelevant version in which relevant items were removed or changed to produce a gap (bad version).

METHOD

Six versions of a logic program were prepared. One of the versions (good version) was modified by the introduction of gaps and irrelevancies and by the removal of branching contingencies, thus producing the other five versions. The six versions were administered to six treatment groups. The resulting posttest scores, adjusted for I.Q. differences, were compared for the six groups.

Sample

Ninety-one paid volunteers who were freshmen and sophomores at local colleges constituted the six treatement groups. The mean age of the students was 18.59 (S.D. = 0.92). The mean I.Q. (Otis, gamma form) was 119.05 (S.D. = 8.52). Students were assigned to one of the six groups in the order in which they appeared at the laboratory.

Materials

The six experimental versions of the logic program are described below.

Good Version. The good version of the logic program was organized into four books having the following content:

Book 1	Introduction to statements Simple and compound statements Letters to symbolize statements
Book 2	Connective "and" Connective "or" Negative "not"
Book 3	Use of multiple connectives Connective "if. . .then. . ." Review Introduction to arguments
Book 4	Valid disjunctive argument Invalid disjunctive argument

This was a branching program; the number of items seen by a student depended on his performance on diagnostic test items (see Appendix B) which are interspersed in the program. The diagnostic test items for Book 4 required the student to:

(1) symbolize a written argument

(2) identify elements in a symbolized argument

(3) determine the validity of a valid argument

(4) determine the validity of an invalid argument.

If the student made an error on any of these test questions (Form A) he was branched to the specific sections in Book 4 covering the topics on which he had made errors. When he completed those sections he was given a parallel form (Form B) of the same diagnostic test. Further errors caused him to be branched back to the appropriate sections again. This cycle was repeated until the student answered all the diagnostic items correctly. He was then given the criterion test.

Linear Version. The linear version was identical to the good version except that the student was not branched back to the program if he made errors on the diagnostic test questions at the end of Book 4. Mastery of the disjunctive arguments was not required of the student prior to his taking the criterion test. Thus the linear version was the good version without remedial branching on the last book.

Small-Gap Version. The small-gap version covered the same topics as the linear version except that some items were modified and others deleted, creating a gap. The deleted items had explained to the student that the truth or falsity of the premises and conclusion of an agrument do not affect its validity and that the inclusive "or" in the first premise disagrees with the second premise and the conclusion in the disjunctive fallacy. The program still provided practice in discriminating between valid and invalid disjunctive arguments, in addition to defining the two argument forms. Thus, the small-gap version was essentially the linear version with a small conceptual gap.

Irrelevant Version. The irrelevant version was identical to the linear version except that two kinds of irrelevant items were added to the program. Items were added which required irrelevant responses to relevant stimuli; these items called for Latin names of the argument forms and covered alternative symbolic equivalents to the logical connectives. Items were also added which covered material not in the posttest; these items covered the construction of truth-tables, but gave no training in their use. Thus, the irrelevant version was the linear version plus irrelevant items.

Small-Gap Bad Version. The small-gap bad version was a combination of the linear, small-gap and irrelevant versions. This version was designed to assess the cumulative effects of the three variables.

Large-Gap Bad Version. The large-gap bad version was identical to the small-gap bad version except that an additional gap was inserted contiguous to the existing small gap. In this version, 34 discrimination practice items were deleted from the small-gap bad version. This version was added to the other five groups to assess the effect of gap size on learning.

A summary of the different versions is shown in Table 2.

Table 2

Description of Treatment Conditions

Version	Branching on Final Book	Gap	Irrelevancies
1. Good	Yes	No	No
2. Linear	No	No	No
3. Small-gap	No	Small	No
4. Irrelevant	No	No	Yes
5. Small-gap bad	No	Small	Yes
6. Large-gap bad	No	Large	Yes

The maximum possible number of program items for each version, (assuming students did not skip any items) is presented in Table 3.

Table 3

Total Number of Items in Each Version

Version	Book				
	1	2	3	4	Total
1. Good	50	52	53	121	276
2. Linear	50	52	53	71	226
3. Small-gap	50	46	53	49	198
4. Irrelevant	50	52	53	108	263
5. Small-gap bad	50	46	53	86	235
6. Large-gap bad	50	46	53	52	201

Criterion Test

The posttest contained 58 questions (Appendix C). Twenty of these covered portions of the programmed material that had not been modified, e.g., logical connectives, introduction to arguments. Thirty-eight covered portions of the program that had been modified. Thus an experimental and a non-experimental subtest score was obtained from each student.

The posttest was substantially the same test used in a previous experiment* that covered the same logic material. Additional items were added to the old test (whose split-half reliability was .97 after Spearman-Brown correction) to fully assess the extent to which students could discriminate between valid and invalid disjunctive arguments.

Procedure

Students came into the SDC CLASS laboratory during the month of May, 1964. They received the Otis Quick-Scoring Mental Ability Gamma Test which took 30 minutes. Directions for using the equipment and the instructional material were then read to them (Appendix A). Students responded to questions in their programmed booklets by pressing buttons on a multiple-key response device. Their answers were recorded by the computer, which also provided them with knowledge of results via small colored lights on the response device.

After the students finished the program version which had been assigned to them, they received a diagnostic test (Appendix B). Individuals in the good group, contingent on their performance on the diagnostic test, were branched back in the program. Members of the other treatment groups were also given the diagnostic test, but did not branch. All students received a posttest at the conclusion of their instruction.

RESULTS AND DISCUSSION

Posttest means and standard deviations adjusted by covariance analysis for I.Q. differences are presented in Table 4. The analysis of covariance for the portion of the test covering the experimentally modified segment of the program was significant ($F = 3.82$, $p < .01$), but no significant difference ($F = .69$, $p > .05$) was obtained among treatment groups on the portion of the posttest covering topics in the program that had not been modified. (The assumption of common regression slopes was tested and confirmed in each case.)

Multiple comparisons were made among the treatment groups on the experimental portion of the test in order to isolate the specific effects of the variables.

---

*H. Silberman, R. Melaragno, J. Coulson, and D. Estavan, "Fixed Sequence vs Branching Autoinstructional Methods," Journal of Educational Psychology, 52:166-172, June, 1962.

Table 4

Adjusted Mean Scores and Significance Tests on Experimental and Non-experimental Segments

Group	N	Experimental Subtest (38 max.) M	SD	Non-experimental Subtest (20 max.) M	SD
1. Good	16	31.17	5.39	15.98	2.84
2. Linear	17	31.48	4.89	16.17	2.61
3. Small-gap	15	27.96	4.16	16.20	2.50
4. Irrelevant	13	27.65	7.45	14.50	4.98
5. Small-gap bad (3 + 4)	17	25.57	5.94	15.67	2.15
6. Large-gap bad (5 + additional gap)	13	21.34	6.71	15.79	2.18
Comparisons					
1 vs. 2		n.s.d.		n.s.d.	
2 vs. 3		1.86*		n.s.d.	
2 vs. 4		1.94*		n.s.d.	
4 vs. 6		3.02**		n.s.d.	

* t-test $p < .05$

** t-test $p < .01$

Results of these comparisons are shown at the bottom of Table 4. The comparison of the first two groups shows no significant difference attributable to branching the student back to the program. The comparison between the second and third groups indicated that the presence of even a small gap in the program had a significant effect on learning. The comparison between the second and fourth groups showed that the addition of irrelevant items also had a significant effect. The comparison between the fourth and sixth groups yielded a significant effect for the large gap. The means in Table 4 also show that the effects of two of the variables are cumulative. That is, the program with only a gap or irrelevant items is superior to the program with both.

In addition to being cumulative, the effects were also specific to the modified portion of the program. Comparisons of criterion test scores covering portions of the program that were the same for the six treatment groups yielded no significant differences. This result would probably not have been obtained if a segment early in the program had been selected for modification. Defective learning on an early program segment would have interfered with subsequent portions of the program. In fact, within the disjunctive argument segment of the program, the small gap placed at the beginning of the segment interfered with learning in later parts of the segment. It was for this reason that the last segment of the program on disjunctive arguments was modified rather than earlier segments. The specificity of the effect of the variables could then be assessed by comparing differences on modified and unmodified parts of the program.

The mean training times and percentage of errors on the program, adjusted for initial differences in I.Q., and the number of items seen in the program are presented in Table 5. The analysis of covariance for the training times was not significant ($F = 2.13$, $p > .05$). This result is interesting in view of the significant differences in the number of items that were seen ($F = 33.47$, $p < .01$) and the percentage of program errors that were made by the six groups ($F = 3.37$, $p < .01$). For example, even though group six had a relatively small number of items (160.54), those items were more difficult (18.58% error) and took longer to complete (71.37 minutes). Training time seems to be more closely related to item difficulty than to the number of items seen.

Table 5

Adjusted Mean Training Times, Adjusted Mean Program Errors, and the Number of Program Items Seen

Group		Training Time (min.)		Percentage of Errors		Number of Items	
	N	M	SD	M	SD	M	SD
1. Good	16	58.43	13.42	12.24	4.65	173.81	26.20
2. Linear	17	61.61	12.54	12.56	8.07	163.00	7.90
3. Small-gap	15	63.44	9.51	16.54	16.07	150.40	9.21
4. Irrelevant	13	69.60	15.08	17.17	10.51	212.77	20.14
5. Small-gap bad (3 + 4)	17	71.09	12.28	20.81	21.71	192.59	13.94
6. Large-gap bad (5 + additional gap)	13	71.37	16.23	18.58	7.30	160.54	13.30

In short, two of the three independent variables, gaps and irrelevancies, had a significant, cumulative, and specific decremental effect on posttest performance. These effects were not obtained at the cost of giving the good groups added training time; if anything, the data suggest that the groups who took the greatest amount of training time received the lowest scores on the portion of the criterion test covering the program segments that had been experimentally modified.

It appears that a program should explicitly include all of the skills, and only those skills, that are relevant to the criterion test. Most instructional material can probably be improved by filling gaps and eliminating irrelevancies. Much work is needed in developing methods of analyzing the relevancy of program content to criterion tests.

While it is possible that the addition of remedial branching does not improve a linear program, as a comparison of the good version and linear version scores would indicate, two alternative explanations are possible. First, it may be that the diagnostic questions used to determine the need for branching did not assess the difficulties students encountered on the posttest. Second, it may be that the remedial items used were not adequate to overcome the students' lack of learning. Earlier research on branching has shown that using alternative remedial materials is often preferable to repeating items in the basic program, as was done in this experiment. From the results of the uses of branching during the tutorial phase of the study, there is evidence that diagnostic test questions and remedial material must be developed carefully, based on extensive empirical trials.

SUMMARY

This research has demonstrated that the exploratory tutorial approach to program evaluation and revision provides a powerful technique for developing effective programmed materials and for generating hypotheses about important program variables.

Three generalizations, or principles, were induced from the observations made during the tutorial sessions in the four subject areas. The principles are: (1) every skill included in the program objectives, and every subskill that logically contributes to those objectives, should be explicitly covered by the program unless it exists in the student's entry repertoire; (2) any materials that do not contribute to the program objectives, as indicated by logical analysis, should be eliminated; and (3) the student should be required to demonstrate mastery of each component subskill before he is allowed to advance to new topics that are based on the earlier materials.

The three principles of program improvement, filling gaps, eliminating irrelevancies, and requiring mastery, although intuitively obvious, bear the added strength of having been induced empirically from independent observations. Two of these principles received experimental verification with another program and a new sample of students.

Other instructional variables of potential importance were also identified, but were not tested in the present study. These include, for example, variables outside the program such as the effect of allowing the student optional control over the instructional sequence, or allowing the teacher to control the pace of the instruction contingent on frequent assessments of the student's behavior. Such "non-program" variables will be studied next in the SDC research program.

(page 44 blank)

APPENDICES

PRECEDING PAGE BLANK-NOT FILMED

APPENDIX A

Experimental Directions

This is an experimental study of the effectiveness of different methods of teaching Symbolic Logic. After you have studied the logic materials, you will be tested on them to determine how effective they were in teaching you the subject.

Here is how to use the equipment and the instructional material:

1. You will be directed to an item in your booklet by a number in the windows of the response device on your desk.

2. Find that item in your booklet and read all the text material on that item very carefully. Study it as you would a textbook.

3. If no answer is required of you, just press the black enter bar when you are ready for the next item. The number in the window will tell you which items comes next. If a question is asked you, examine it completely and select an answer. Then find the button with the letter of your choice on your response device and press it all the way down. If you decide to change your answer, just press another button. When you want the machine to evaluate your answer, press the enter bar and a green light will come on if you got it right; a red light if you got it wrong. Another light will be turned on opposite the correct answer.

4. If all of the lights go on at once, raise your hand for help. That probably means you are ready to go on to the next booklet whose number will appear on your windows.

5. When you are ready to go on to the next item press the enter bar again and a number will appear on the little windows of your response device telling you what item in your booklet to study next. Always be sure you are looking only at the item whose number appears on your window. This cycle repeats itself with each new item. (Incidentally, you will find the items in your booklet are numbered consecutively on the top row, then when you reach the last page, you will return to the first page and the items continue on the bottom row of items.

6. The questions are written to be easy to answer and you shouldn't have difficulty. If you make many errors, you will know that you haven't been studying as carefully as you should.

7. Don't ask for assistance with the subject matter. All the help you need will be found in the booklet.

8. Make no marks in the booklet. You are encouraged to take notes on your scratch pad as this will help you do better on the test.

9. In summary,

(a) Study the item in your booklet whose number appears in the windows of your response device.

(b) Insert an answer if required and press the enter bar. Then press the enter bar again when you are ready to go to the next item. If no answer is required,

(c) Just press the enter bar when you are ready for the next item.

(d) Turn to the new item in your booklet whose number appears on your response device and repeat the cycle. Remember only read the item whose number appears on your window.

Open books to page 1 and begin.

APPENDIX B

Diagnostic Test - Form A

1.

Either we will not go to Rome or we will not go to Naples.
Since we're going to Naples, we won't go to Rome.

Symbolize this argument:

2.

$$\begin{array}{l} b \vee c \\ b \\ \hline \therefore \sim c \end{array}$$

Determine the validity of this argument: _____

3.

$$\begin{array}{l} r \vee p \\ \sim p \\ \hline \therefore r \end{array}$$

Determine the validity of this argument: _____

4.

$$\begin{array}{l} f \vee \sim g \\ \sim g \\ \hline \therefore \sim f \end{array}$$

In this argument, the first premise is a compound statement which has a connective called the _____.

Does the second premise deny one of the simple statements in the first premise? (Yes, No) _____.

Does the conclusion repeat one of the simple statements in the first premise? (Yes, No) _____.

5.

$$\frac{\sim h \vee \sim k \quad h}{\therefore \sim k}$$

Determine the validity of this argument:_____

6.

$$\frac{p \vee \sim q \quad \sim q}{\therefore \sim p}$$

Determine the validity of this argument:_____

Diagnostic Test - Form B

1.

Either dried grapes are not raisins or they are not pomegranates. Since dried grapes are raisins, they are not pomegranates.

Symbolize this argument:

2.

$$\frac{x \vee z \quad x}{\therefore \sim z}$$

Determine the validity of this agrument:_____

3.

$$\begin{array}{l} d \vee t \\ \sim t \\ \hline \therefore d \end{array}$$

Determine the validity of this argument: _____

4.

$$\begin{array}{l} r \vee \sim s \\ \sim s \\ \hline \therefore \sim r \end{array}$$

In this argument, the first premise is a compound statement which has a connective called the _____.

Does the second premise deny one of the simple statements in the first premise? (Yes, No) _____.

Does the conclusion repeat one of the simple statements in the first premise? (Yes, No) _____.

5.

$$\begin{array}{l} \sim p \vee \sim q \\ p \\ \hline \therefore \sim q \end{array}$$

Determine the validity of this argument: _____

6.

$$\begin{array}{l} c \vee \sim d \\ \sim d \\ \hline \therefore \sim c \end{array}$$

Determine the validity of this argument: _____

APPENDIX C

Logic Criterion Test

Mark an X on the line before the answer of your choice.

For example:

What is the subject you were just taught?

____ 1. mathematics
_X_ 2. symbolic logic
____ 3. algebra
____ 4. history

1. How many simple statements are there in a compound statement?

____ 1. None
____ 2. One
____ 3. Two
____ 4. Two or more

2. Choose the simple statements (more than one).

____ 1. Turtles lay eggs.
____ 2. The bright summer sun.
____ 3. Is Clark Gable 60 years old yet?
____ 4. The quick brown fox jumped over the lazy dog.
____ 5. Shake the can.

3. One way of distinguishing a statement from any other type of sentence is that a statement:

____ 1. has a predicate and an object.
____ 2. has at least one dependent and one independent clause.
____ 3. must be either true or false.
____ 4. can be either simple or compound.

4. In logic _____ are used to represent statements.

_____ 1. letters
_____ 2. words
_____ 3. figures
_____ 4. connectives
_____ 5. symbols

5. A connective is used to:

_____ 1. combine the parts of a simple statement.
_____ 2. combine statements into an argument.
_____ 3. connect the members of the object class.
_____ 4. combine simple statements into compound statements.

6. Which of the following does <u>not</u> belong with the others?

_____ 1. and
_____ 2. or
_____ 3. but
_____ 4. although
_____ 5. however

7. What is meant by $p \wedge q$?

_____ 1. p but also q
_____ 2. p but not q
_____ 3. q but not p
_____ 4. neither p nor q
_____ 5. p is greater than q

8. In logic only the _____ "or" is used.

_____ 1. inclusive
_____ 2. exclusive
_____ 3. conditional
_____ 4. conjunctive
_____ 5. disjunctive

9. Which of the following is true of the disjunction?

____ 1. The symbol is $\vee$ --the translation is "or."
____ 2. The symbol is $\wedge$ --the translation is "and."
____ 3. The symbol is $\wedge$ --the translation is "or."
____ 4. The symbol is $\vee$ --the translation is "and."

10. The inclusive use of the statement "l $\vee$ w" means:

____ 1. l or w, or both.
____ 2. both l and w.
____ 3. l or w, but not both.
____ 4. when l is true w is true.

11. Let "p" stand for "acid dissolves tin," and "q" stand for "acid turns copper black." You are given on absolutely reliable authority the statement "p $\vee$ q." In an experiment you find that acid does dissolve tin. Your next step should be to:

____ 1. do a second experiment to determine whether acid turns copper black.
____ 2. do the first experiment over, because the tin should not have dissolved.
____ 3. stop, because you now know the acid must turn copper black.
____ 4. stop, because you now know the acid will not turn copper black.

12. What is the negation of $\sim$ (glab $\vee$ %)

____ 1. glab $\vee$ %
____ 2. $\sim$ glab $\vee$ %
____ 3. glab $\wedge \sim$ %
____ 4. $\sim$ glab $\vee \sim$ %

13. Let w stand for "I am warm," and let c stand for, "I will turn on the cooler." What does "c $\longrightarrow$ w" mean?

____ 1. "If I turn on the cooler I am warm."
____ 2. "If I am warm I turn on the cooler."
____ 3. "I will turn on the cooler or I am warm."
____ 4. "I am warm and I turn on the cooler."
____ 5. "I am warm but I will not turn on the cooler."

14. The conditional of "p" and "$\sim$ q" is:

____ 1. $p \wedge \sim q$
____ 2. $P \rightarrow \sim q$
____ 3. $p \rightarrow q$
____ 4. $\sim p \leftrightarrow q$

15. What is the symbolic representation for the sentence "The rain was heavy, but it did not last long?"

____ 1. $p \vee \sim q$
____ 2. $\sim p \rightarrow q$
____ 3. $p \wedge \sim q$
____ 4. $p \wedge q$
____ 5. $\sim p \wedge \sim q$

16. The disjunction of "$\sim$ p" and "$\sim$ q" is:

____ 1. $\sim p \wedge \sim q$
____ 2. $\sim p \vee q$
____ 3. $\sim p \rightarrow \sim q$
____ 4. $\sim p \vee \sim q$

17. Which connectives are used in symbolizing the following statement?

"Some students do not know about computers although they are becoming more important."

____ 1. negation, conjunction
____ 2. conditional, disjunction
____ 3. disjunction, negation
____ 4. conjunction, conditional
____ 5. negation

18. What is the symbolic form for the statement, "It is not true that either Johnson or Nixon will win, and it is also not true that Mao Tse Tung is running."

____ 1. $\sim(k \vee n) \wedge \sim m$
____ 2. $(k \vee n) \wedge \sim m$
____ 3. $(k \vee n) \wedge m$
____ 4. $\sim(k \vee n) \vee \sim m$
____ 5. $k \vee n \vee m$

Write in your answers to the following questions on the blank lines.

For example:

The subject you were taught is __symbolic__ __logic__.

19. Let p represent the statement "Thinking is hard work," and let q represent the proposition, "I was a fool to take this course." Write out in words the meaning of each of the following symbols:

a. $p \wedge \sim q$

20. For symbolizing the statement "American Airlines uses Boeing jets or Douglas jets," one could use "p" for the first simple statement, "q" for the second simple statement, and _____ for the connective.

21. In determining the validity of an argument remember that:

____ 1. if the conclusion is true, the argument is valid.
____ 2. if the conclusion is false, the argument is valid.
____ 3. the form of the argument determines its validity.
____ 4. unless the premises are true, the argument is a fallacy.

22. What two things do the following arguments have in common?
(Place check marks in front of your choices below)

$$\begin{array}{l} p \vee q \\ \underline{p} \\ \therefore \sim q \end{array} \qquad\qquad \begin{array}{l} p \vee q \\ \underline{q} \\ \therefore \sim p \end{array}$$

_____ 1. They both contain the disjunction connective.
_____ 2. They both contain the conjunction connective.
_____ 3. They both contain the conditional connective.
_____ 4. They are both valid.
_____ 5. They are both fallacies.

23. Using the letters "p" and "q" symbolize the following argument:

Carol is rich or happy.
She is not happy.
Therefore, she must be rich.

24. Determine the validity of the following argument:

Carol is rich or happy.
She is not happy.
Therefore, she must be rich.

_____ 1. Valid.
_____ 2. Invalid.
_____ 3. This is not an argument.
_____ 4. Not enough information is given.

25. Using the letters "p" and "q" symbolize the following argument:

Either I do well in sports or I cannot be proud of myself.
I do well in sports.
Therefore, I am proud of myself.

26. Determine the validity of the following argument:

Either I do well in sports or I cannot be proud of myself.
I do well in sports
Therefore, I am proud of myself.

_____ 1. Valid.
_____ 2. Invalid.
_____ 3. This is not an argument.
_____ 4. Not enough information is given.

Symbolize each of the following arguments and indicate whether it is valid or invalid.

27-28. The belt is not too small or I'm not too fat.
The belt is not too small, therefore I am too fat.

29-30. Either I am not a beatle or I don't have long hair.
I have long hair so I am not a beatle.

31-32. Wasps have either stingers or no antenna.
Wasps have no stingers, therefore they don't have antenna.

33-34. Elvis sings or he plays the guitar
Elvis does not play the guitar, therefore he sings.

35-36. Either the barometer is falling or there will be a storm.
The barometer is falling, so there will not be a storm.

37-38. One must learn Latin well or one will never understand English.
You learned Latin well, therefore you understand English.

39-40. Either the carburetor is not flooded or the car will run.
The car doesn't run, therefore the carburetor is not flooded.

41-42. Snakes cannot talk or they would be able to communicate.
Snakes do communicate, therefore they can talk.

43-44. The children will not go to bed or they will watch television.
The children will not go to bed, therefore they won't watch television.

45-46. Tito is not a capitalist or Yugoslavia would not hold free elections.
Yugoslavia does not hold free elections so Tito is a capitalist.

47-48. Either you or I didn't hang up the telephone after our conversation last night.
I hung up the telephone, therefore you didn't.

49-50. I will break that record or my name is John Smith.
My name is John Smith so I will not break that record.

51-52. No man lives on Mars or all men would not speak Martian.
A man lives on Mars, therefore all men don't speak Martian.

53-54. Either he didn't recognize me or he was snubbing me.
He did recognize me so he was snubbing me.

55-56. Either hail or rocks are hard.
Hail is hard, therefore rocks are not hard.

57. To determine the validity of an argument first notice whether or not:

1. the premises are true statements.
2. the conclusion is related to the first premise.
3. the second premise agrees with part of the first premise.
4. the conclusion is a true statement.

58. The disjunctive fallacy is invalid because:

1. it does not contain a compound statement.
2. the truth values for the first premise disagree with the truth values for the conclusion.
3. the inclusive "or" in the first premise disagrees with the second premise and the conclusion.
4. the conclusion is false.

E N D. DEPT. OF HEALTH EDUCATION & WELFARE U.S. OFFICE OF EDUCATION E R I C.
DATE FILMED 6. 26. 67